DOCUMENT RESUME

ED 121 846 TM 005 278

AUTHOR Gomes, Marcia P. R. de Magalhaes; And Others
TITLE Validation of the Teacher's Assessment of Students' Aptitudes Form.
PUB DATE [Apr 76]
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS *Aptitude Tests; *Classroom Observation Techniques; Grade 8; Junior High Schools; Lower Middle Class; Statistical Analysis; *Student Behavior; Student Evaluation; *Test Validity; *Vocational Aptitude
IDENTIFIERS Battery of Differential Aptitudes Tests; Brazil (Rio de Janeiro); *Teachers Assessment of Students Aptitudes Form

ABSTRACT

The purpose of this study was to validate the Teacher's Assessment of Students' Aptitudes Form (TASAF), an instrument devised to guide teachers assessment of students' aptitudes through the observation of their classroom behaviors, i.e., to what extent were teachers able to detect levels of differential aptitudes in their students through guided observation. Since the TASAF was designed to tap the same set of aptitudes included in the Battery of Differential Aptitudes Test (BDAT), the BDAT was used as the criterion for the validation of the TASAF. Students enrolled in the eighth grade of 9 Brazilian public school located in Rio de Janeiro took part in the study. The group was composed of girls and boys ranging in age from 13 to 18 years. The subjects belonged mostly to the middle-lower class, as indicated by their parents' occupation. Data treated by the "G" index of agreement indicated statistically significant correlations between measures on the two instruments. Results pointed to a promising alternative approach to the traditional method of assessing students' aptitudes through paper-pencil tests. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

Session
H 6.12

ED121846

VALIDATION OF THE TEACHER'S ASSESSMENT OF STUDENTS' APTITUDES FORM

by

Marcia P.R. de Magalhães Gomes
Lilia da Rocha Bastos
Ruth Nobre Scheeffer

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

1976 ANNUAL MEETING

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

SAN FRANCISCO, CALIFORNIA

TM005 278

6.12

VALIDATION OF THE TEACHER'S ASSESSMENT OF STUDENTS' APTITUDES FORM -- TASAF

MARCIA P.R. MAGALHÃES GOMES
Federal University of Rio de Janeiro

LILIA DA ROCHA BASTOS
Federal University of Rio de Janeiro

RUTH NOBRE SCHEEFFER
Federal University of Rio de Janeiro

The Brazilian Educational Reform Law enacted in 1971 (Brasil, Congresso Nacional, 1971) has prescribed the introduction of vocational education into the country's system of education starting in grade 9, or even earlier, whenever local and/or individual characteristics should recommend it. The Law has also required that the process of assessing students' aptitudes should begin as soon as they enroll in the elementary school and go on until the 8th grade, with greater emphasis placed on the period between the 5th and 8th grades. One year later, the Federal Council of Education (Conselho Federal de Educação, 1972) indicated that the process of assessing students' aptitudes should be jointly undertaken by teachers, school guidance workers and parents and use as main source of data the continuous observation of the student behavior in different school activities.

The lack of measuring instruments to guide teachers in the process of continuously assessing students' aptitudes, the insufficient number and high costs involved in the administration of standardized aptitude tests in Brazil have given rise to the problem of this study.

Objective

It was the purpose of this study to validate the Teacher's Assessment of Students' Aptitudes Form -- TASAF, an instrument devised to guide teachers assessment of students' aptitudes through the observation of their classroom behaviors. In other words, its objective was to investigate to what extent were teachers able to detect levels of differential aptitudes in their students through guided observation.

Since the TASAF was designed to tap the same set of aptitudes included in the Battery of Differential Aptitudes Tests -- BTAD (Fundação Getúlio Vargas, 1975), the BTAD was used as the criterion for the validation of the TASAF.

Assumptions

The study was based on two assumptions:

1. Students' behaviors in school situations reflect a representative range of differential aptitudes.

2. Teachers are able to identify students' differential aptitudes through guided observation of their classroom behaviors.

Hypotheses

The research hypothesis stated that teacher's assessment of students' aptitudes through the observation of their classroom behaviors would agree with measures of aptitudes obtained through the administration of the Battery of Differential Aptitudes Tests (BTAD).

Twenty-two null hypotheses, stating no significant agreement between students' atptitudes as measured by the BTAD and observed by the teachers with the use of the TASAF, were derived from the research hypothesis. Each null hypothesis referred to one of the 11 tests of the BTAD and their corresponding TASAF items and to the subjects' sex.

Methods

Subjects

Students enrolled in the 8th grade of a Brazilian public school located in Rio de Janeiro took part in the study. The group was composed of 100 girls and 120 boys whose ages ranged from 13yrs. to 18yrs. The subjects belonged mostly to the middle-lower class, as indicated by their parents' occupation.

The teachers who rated the students' aptitudes totalled 26 (Table 1).

Table 1

Distribution of Teachers by Discipline

Discipline	Number
Arts	4
Math	3
Physical Education	4
Portuguese	2
Science	3
Social Studies	3
Business	2
Industrial Arts	2
Home Economics	3
T O T A L	26

Instrumentation

Two instruments were used in this study: the Battery of Differential Aptitudes Tests -- BTAD, Form 1.A (Fundação Getúlio Vargas, 1975) and the Teacher's Assessment of Students' Aptitudes Form -- TASAF.

The BTAD was developed by a group of Brazilian and American psychologists under the supervision of the American Institutes for Research -- AIR -- and the sponsorship of the Agency for International Development and the Brazilian Government. The Battery includes 11 tests: Vocabulary, Problems, Verbal Analogies, Numerical Series, General Information, Arithmetic, Tables, Names, Figures, Mechanical Knowledge, and Space Relations.

The BTAD manual reports that the concurrent validity coefficients of correlation for most of its tests, taking as criteria 5th graders' achievement scores in disciplines related to the content of the tests, were found to be significant at .05.

The coefficients of reliability of the tests calculated both by the test-retest method and by the KR-20 formula were above .70 with the exception of the General Information test whose reliability was .59.

As to the interpretation of the tests'scores for the purpose of vocational guidance, the BTAD manual indicates that individuals scoring at the 50th percentile or below in factors judged to be essential for the performance of specific occupations should not be considered fitted those occupations while those scoring above the 50th percentile should. The BTAD reports norms for 5th, 6th, and 7th graders. Although the Ss of this study were 8th graders norms for 7th graders were used since the Battery was administered to the students at the outset of the school year, when their knowledge would be that of 7th graders.

The Teacher's Assessment of Students' Aptitudes Form -- TASAF (Appendix 1) is an instrument developed for the teacher to assess students' aptitudes through the observation of their behavior in regular school situations. The TASAF includes 11 items to be evaluated by the teacher through the observation of his students' behaviors. Each TASAF item corresponds to a factor or aptitude tapped by a BTAD test. For example, the TASAF item Knowledge of Vocabulary corresponds to the BTAD Vocabulary test.

Each TASAF itemwas operationally defined in terms of behavior indicators in order to guide the teacher's observation of each student aptitude and to help minimize teacher's subjectivity. For instance, among the examples of students' behaviors that can be observed in the assessment of Space Relations, the following were suggested by the TASAF: use of blackboard and paper space, movements and actions performed during physical exercises and games, and placement of given elements into three dimensional spaces using figures as guides.

Using and Rating the TASAF Items

While some of the TASAF items involve general behaviors that can be observed by most teachers irrespective of the disciplines they teach, other items involve specific behaviors related to certain disciplines. For example, Knowledge of Vocabulary is an item which can be illustrated by the following behaviors: use of common and uncommon words of both general and specialized usage according to students' grade level and age.

Vocabulary of general usage can be observed by any teacher through students' oral or written communication. Students' use of specialized words can be also rated by any teacher through the observation ot their use of words related to his own discipline. Arithmetic Reasoning, on the other hand, being an item indicated by the ability to perform numerical operations, to solve numerical problems and to handle numerical concepts, can be systematically observed by Math and Science teachers and only eventually by other teachers.

The TASAF manual offers indicators of behaviors to be observed for each item of the Form (Appendix 2). Teachers are warned not to rate any item whose behavior indicators they were not able to observe in their students.

Student behaviors for each TASAF item are rated in a five-letter ordinal scale similar to the one used by teachers to rate students' academic achievement: A, B, C, D, and E. The first three letters (A, B, and C) are considered passing grades and mean that more than 50% of the criterion was attained. The last two (D and E) mean an achievement equal to or below 50% of the criterion. In order that the scoring of the TASAF were congruent with that of the BTAD, that is, in order that it were also a norm-referenced scoring, teachers were asked to rate their students in comparison with all the students they had ever taught at the same grade and age level. Thus, students who received ratings A, B, or C could be considered superior to 50% of their peers of similar grade and age levels (Appendix 3). Such a procedure allowed the researchers to dichotomize the five-letter scale into:(a) A, B and C, meaning positions superior to P_{50}; and (b) D and E, meaning positions equal to or below P_{50}.

Data Collection

The BTAD was administered to the subjects in the beginning of the 1975 Spring(1) Semester by Psychology majors of a local university under the coordination of a staff member of the Institute for Selection and Professional Guidance (ISOP), in Rio de Janeiro.

The TASAF items were filled out for each student by each one of his teachers, separately. The final rating for each student was decided upon by all his teachers together in what is called a class council, in the same fashion that students' academic achievement is regularly evaluated. The school guidance worker assisted the teachers in the use of the TASAF during the semester the study was undertaken.

Statistical Treatment

Students' scores in the BTAD were correlated with students' ratings obtained in the TASAF through the "G" index of agreement (Lienert, 1972). Scores for each student in the BTAD were dichotomized into above and equal to or below the 50th percentile,

---

[1] Autumn Semester in Brazil

following the BTAD norms for 7th graders. Students' ratings in the TASAF were also dichotomized into above P_{50}, encompassing ratings A, B, and C; and equal to or below P_{50}, for ratings D and E, for each item.

Null hypotheses of no significant agreement between students' results in the BTAD and in the TASAF were tested for each aptitude under assessment, by a procedure suggested by Lienert (1972). The level of significance for all the tests performed was set at .05.

Results

With one exception, represented by the lack of agreement between the male subjects' scores in the Figures test of the BTAD and its corresponding TASAF item, all the null hypotheses tested in the present study were rejected. The "G" indices of agreement and their respective critical ratios are presented in Table 2.

Table 2

"G" Indices of Agreement Between Students' Results on the BTAD and on the TASAF and their Corresponding Critical Ratios, by Aptitude and Sex

Aptitude	"G" Index of Agreement		Critical Ratio	
	Boys	Girls	Boys	Girls
Vocabulary	.65	.70	6,82	7,63
Problems	.54	.68	5,64	7,38
Verbal Analogies	.82	.73	8,60	7,93
Numerical Series	.72	.81	7,16	8,72
General Information	.72	.50	6,94	5,39
Arithmetic	.24	.43	2,40	4,63
Tables	.57	.53	5,61	5,71
Names	.61	.67	6,04	7,22
Figures	.18*	.40	1,80	4,30
Mechanical Konowledge	.72	.62	7,20	6,67
Space Relations	.57	.65	5,87	7,18

(*) Only index of agreement not statistically significant at .05

The results seem to indicate that the TASAF enabled teachers to assess students' differential aptitudes at a reasonable degree of accuracy, taking the BTAD as criterion. The lack of significance observed in the Figures for the sample of boys could be tentatively explained by the difficulty inherent at the observation of students' ability to perceive figures, by the fact that boys in general participate less overtly than do girls in class activities, or by both factors.

The teachers were generally highly cooperative and seemed to be anxious to adequately rate their students' aptitudes.

During the class council, they demonstrated that did not use the students' academic achievement to rate their aptitudes. In fact, they often rated students' aptitudes high, even though the students did not evidence a high achievement in areas correlated to the aptitude, emphasizing that they were not evaluating achievement but rather aptitude. These observations indicate that the TASAF might yield less culture loaded assessments of students' aptitudes than the regular aptitude tests do. It was also observed that the teachers tended to be somewhat lenient in rating their students' aptitudes fearing that their judgements could somehow jeopardize the students' future professional choice.

Math, Portuguese, Social Studies, Arts, and Science teachers were the most productive in assessing the students' aptitudes. Business, Industrial Arts and Home Economics teachers contributed little and seemed to find it hard to rate their students' aptitudes. This fact may be due to deficiencies in their professional preparation, especially in the areas of Educational Foundations and Evaluation. Physical Education teachers were the least responsive to the study, neither filling out any item of the TASAF nor taking part in the class council discussions.

Conclusions

An evaluation of the results of this study, taking into consideration the conditions under which it was undertaken, indicated the following conclusions:

1. If only one test of the BTAD -- Figures, in the sample of boys -- did not correlate significantly with its TASAF corresponding item, it can be inferred that teachers are able to adequately assess their students' aptitudes through the use of the TASAF.

2. If the TASAF worked out when used in the 8th grade, it is fair to expect similar results when the Form is used in other grade levels.

3. If the TASAF yielded positive results when used during only one semester, more accurate results may occur when the use of the Form is extended over a year.

4. If future studies indicate results similar to those herein obtained, the application of the TASAF might be generalized to the country's system of education since (a) it does not involve costs other than those of buying one Form for each teacher/class and one instructional booklet for each teacher; and (b) it does not require specialized personnel to its administration as aptitude tests do, in Brazil.

REFERENCES

Brasil, Congresso Nacional. Lei 5692/71, Diário Oficial, 12 de agosto de 1971.

Conselho Federal de Educação. Parecer 339, Documenta 137, 1972

Fundação Getúlio Vargas. Bateria de Testes de Aptidões Diferenciadas -- BTAD-1.A. Rio: Fundação Getúlio Vargas, 1975.

Lienert, G.A. Notes on tests concerning the "G" index of agreement. Educational and Psychological Measurement, 1972, 32, 281-288.

APPENDIX 1

TASAF	SCHOOL _____ GUIDANCE WORKER _____	YEAR _____ GRADE ____ CLASS ____ DISCIPLINE _____

STUDENTS		RATINGS (A, B, C, D, E)									
Nº	NAMES	Knowledge of Vocabulary	Arithmetic Reasoning	Verbal Reasoning	General Reasoning	Speed and Accuracy in Computation	Interpretation of Graphic Data	Speed and Accuracy in Discrimination	Perception of Pictorial Stimuli	Knowledge of Mechanical Instruments	Visualization of Space Relations

APPENDIX 2

EXAMPLES OF BEHAVIORS TO BE OBSERVED FOR EACH TASAF ITEM

ITEM	BEHAVIORS
Knowledge of Vocabulary	- uses adequately common and uncommon words of both general and specialized usage according to his own age and grade level
Arithmetic Reasoning	- performs numerical computations - solves numerical problems - deals with numerical concepts
Verbal Reasoning	- responds appropriately to written or oral communication - express ideas in oral or written form - identifies absurdities in stories - describes situations - recognizes the central idea in a passage
General Reasoning	- responds by words, gestures or acts to requests - establishes the relationship between facts either in written or oral form - employs this relationship to arrive to new ideas - arrives to conclusions - gives examples of his own - synthetizes - solves problems

ITEM	BEHAVIORS
General Knowledge	- responds orally or in written form to questions referring to: (a) relevant matters of different disciplines; (b) general subjects learned through different media (newspaper, books, TV, movies, etc.)
Spedd and Accuracy in Computation	- performs simple operations (addition, subtraction, multiplication and division) mentally, with accuracy and speed
Interpretation of Graphic Data	- extracts data from tables - interprets maps - interprets scales
Speed and Accuracy in Discrimination	- distinguishes similarities and differences in objects or symbols, with accuracy and speed
Perception of Pictorial Stimuli	- identifies grapnic representations: graphs, symbols, diagrams - recognizes relationships within a graphic representation and between graphs, symbols and diagrams
Knowledge of Mechanical Instruments	- Identifies equipment, devices, tools, and appliances - describes how machines work - indicates how to take care of machines to keep them in order
Visualizations of Space Relations	- uses the blackboard and the paper spaces harmoniously - moves and acts with balance during physical exercises and games - places given elements into three dimensional spaces using figures as guides.

APPENDIX 3

GENERAL INSTRUCTIONS FOR FILLING CUT THE TASAF ITEMS

Start filling out the TASAF after you have had the opportunity to observe those students' behaviors indicated for each item of the Form. As much as possible, restrict your observation to the students' aptitudes, trying not to get influenced by their school achievement in correlate areas. Your regular contact with the class will allow you to distinguish those students who, in spite of demonstrating specific aptitudes, do not achieve high in correlate areas, from those with real learning problems in these areas.

Leave it blank any item whose behaviors you did not have the opportunity to observe. Probably, other teachers will have the chance to observe them on account of the nature of the disciplines they teach.

The position of the student in each TASAF item will be expressed by the letter ratings A, B, C, D, and E, which mean, respectively, an EXCELLENT, GOOD, FAIR, POOR and DEFFICIENT level in a specific aptitude. In the process of rating a student's aptitude, take as point of reference all the students of similar grade and age level you have ever taught or had contact with.

To rate your students' aptitudes, use the following criteria:

A = students above 90% of their peers of similar age and grade level in the aptitude under assessment.

B = students above 70% up to 90% of their peers of similar grade and age level in the aptitude under assessment.

C = students above 50% up to 70% of their peers of similar grade and age level in the aptitude under assessment.

D = students above 30% up to 50% of their peers of similar grade and age level in the aptitude under assessment.

E = students equal to or below 30% of their peers of similar age and grade level in the aptitude under assessment.